MULTIMEDIA UNIVERSITY®

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2017/2018

THP7021 – HIGH PERFORMANCE COMPUTING AND BIG DATA

(All sections / Groups)

4th October 2017
2.30 p.m - 4.30 p.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This Question paper consists of 4 pages with 4 Questions only.
2. Attempt **ALL FOUR (4)** questions. All questions carry equal marks and the distribution of the marks for each question is given.
3. Please print all your answers in the Answer Booklet provided.

Question 1 [10 marks]

This serial program in *Listing 1*, if asked to calculate a picture of size 1024, will produce a picture like what is shown in *Figure 1*.

```
 1 #include <stdio.h>
 2 #include <stdlib.h>
 3 #include <math.h>
 4
 5 int main(int argc, char*argv[])
 6 { int xmax, ymax, x, y, size;
 7   double divisor;
 8   int **picture;
 9   char buff[1024];
10   FILE *outfile;
11
12   printf("What file should I store the picture in? ");
13   gets(buff);
14   printf("How big should the picture be? ");
15   scanf("%d", &size);
16
17   divisor = (M_PI*size/60);
18   outfile=fopen(buff,"w");
19   fprintf(outfile,"P2\n%d %d 256\n",size, size);
20
21   picture =(int**)malloc(sizeof(int*)*size);
22   for(y=0; y<size; y++)
23   { picture[y] = (int*)malloc(sizeof(int)*size);
24     for (x=0; x<size; x++)
25     { picture[y][x] = (sin(x/divisor)+cos(y/divisor))*64+128;
26       fprintf(outfile,"%d\n", picture[y][x]);
27     }
28   }
29   fclose(outfile);
30 }
```

Listing 1

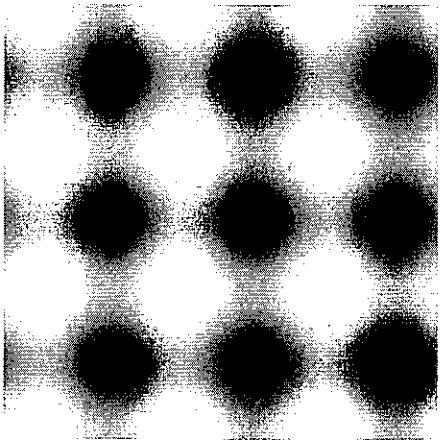

Figure 1

Rewrite the program so that it will be in MPI. Only the master should ask the user for input, and only the master should save the output picture. Beware that you must save the output picture in the right order – so be careful how you gather the data back from the slaves.

Continued...

You must split up the work evenly among all the processors, and you must not hard-code the number of processors.

Refer to the line numbers from *Listing 1* in order to not have to recopy everything, e.g. "// copy L1's lines 7 – 12 here" [10 marks]

Question 2 [10 marks]

Convert the serial program from *Listing 1* into an OpenMP program. Be careful about the inner loop, and the writing of the output file. [10 marks]

Question 3 [10 marks]

a) What is the difference between Map and Reduce in Hadoop? [4 marks]

b) If you have a dataset that contains:
- Product Code
- Product Name
- Supplier Code
- Price

Write an algorithm for the map and reduce to calculate the average price for each supplier. [6 marks]

Continued...

Question 4 [10 marks]

Given this serial programme:

```
double complexCalc(double x); // This is a function that
// will take a lot of computing to calculate.

int main(int argc, char *argv[])
{ double total = 0, x;
  int partitions;
  double slice, start, end;

  printf("How many partitions? "); fflush(stdout);
  scanf("%d", &partitions);
  printf(("Start and end point of integration: ");
  fflush(stdout);
  scanf("%f %f"; &start, &end);

  slice = (end-start)/partitions;
  for (x = start + (slice/2); x < end; x = x + slice)
    total = total + f(x);
  total = total * slice;

  printf("The integration is %1.20f\n", total);
}
```

a) When would it be better to convert this into an MPI program instead of an OpenMP, Pthreads, or MapReduce program? Why would this situation be better for MPI and not so good for OpenMP, Pthreads or MapReduce? [5 marks]

b) When would it be better to convert this into an OpenMP program instead of an MPI, Pthreads, or MapReduce program? Why would this situation be better for OpenMP and not so good for OpenMP, Pthreads or MapReduce? [5 marks]

End of Paper